

مدل هیبریدی فراکاوشی یادگیری ماشینی تخمین سطح آب زیرزمینی

شیوا خسروی^۱، امیر رباطی^{۲*}

۱. دانشجوی دکتری عمران- مهندسی و مدیریت منابع آب، واحد کرمان، دانشگاه آزاد اسلامی، کرمان، ایران

۲. استادیار گروه مهندسی عمران، واحد کرمان، دانشگاه آزاد اسلامی، کرمان، ایران

*A_robati@iauk.ac.ir

تاریخ دریافت: ۹۹/۰۹/۱۸ تاریخ پذیرش: ۹۹/۱۲/۲۸

چکیده

آب چاه به عنوان منبع محدود طبیعی تامین آب، نقش حیاتی در مناطق خشک و نیمه خشک ایفا می‌کند. در سال‌های اخیر با توجه به مشکل کمبود منابع آبی، مساله استفاده و مدیریت بهینه از اهمیت ویژه‌ای برخوردار است. به منظور آگاهی از وضعیت این منابع و مدیریت بهینه آن، پیش‌بینی نوسانات سطح آب چاه لازم است. مدل‌های دقیق و قابل اطمینان پیش‌بینی سطح آب چاه‌ها می‌توانند به استفاده پایدار از آب زیرزمینی به منظور تامین نیازهای شهری، کشاورزی و صنعتی کمک کنند. مدل‌های عددی نیازمند پارامترهای زیاد، زمان بر و پرهزینه هستند، پس در سال‌های اخیر توجه خاصی به مدل هوشمند شده است. در این پژوهش مدل حداقل مربعات ماشین بردار پشتیبان وزن‌دار (WLS-SVM) با دقیق‌ترین تابع وزنی Myriad با الگوریتم‌های فراکاوشی جامعه پرندگان (PSO) و گرانشی (GSA) ترکیب شده است و این دو مدل هیبریدی (WLSSVM-PSO) و (WLSSVM-GSA) به عنوان روش‌های محاسباتی هوشمند جدید به منظور تخمین سطح آب زیرزمینی ارائه شده است. به منظور تخمین سطح آب زیرزمینی از داده‌های ده چاه مشاهداتی در دشت باغین استان کرمان، دارای آمار تراز سطح ایستابی با سری زمانی ده ساله استفاده شده است. مقادیر تخمین زده شده از دو مدل هیبریدی با مقادیر مشاهده‌ای تراز سطح ایستابی مقایسه شد و عملکرد این مدل‌ها با شاخص‌های آماری ارزیابی شده است. نتایج نشان داد که مدل‌های هیبریدی WLSSVM-PSO و WLSSVM-GSA دارای دقت بسیار خوبی برای تخمین سطح آب زیرزمینی است، اما مدل WLSSVM-GSA نسبت به مدل WLSSVM-PSO کمی با دقت بالاتر عمل کرده است. بنابراین این مطالعه نشان می‌دهد که می‌توان از این مدل‌های هیبریدی ارائه شده به عنوان ابزاری کارآمد در تخمین سطح آب زیرزمینی استفاده نمود.

واژگان کلیدی: سطح آب چاه، مدل هیبریدی هوشمند، حداقل مربعات ماشین بردار پشتیبان وزن دار، الگوریتم فراکاوشی.

۱- مقدمه

دخیره سازی بالاتری دارند. علاوه بر این در حالت طبیعی دارای کیفیت بهتری هستند و کمتر در معرض آلودگی هستند. همچنین تلفات این آب‌ها بر اثر تبخیر کمتر بوده و در برابر

آب‌های زیرزمینی بخش عمده آب‌های شیرین در دسترس بشر محسوب می‌شوند. آب‌های زیرزمینی در مقایسه با آب‌های سطحی دارای گستردگی و ماندگاری بیشتری بوده و قابلیت

فیزیکی شده است که از آن جمله می‌توان به مدل‌های هوشمند اشاره کرد. در سال‌های اخیر مدل‌های هوشمند به عنوان ابزار قدرتمند در پیش‌بینی رویدادهای هیدرولوژیکی از جمله سطح آب چاه‌ها پیشرفت چشمگیری داشته است.

روش ماشین بردار پشتیبان یکی از مدل‌های هوشمند یادگیری ماشینی است که می‌تواند روابط غیرخطی و پیچیده را مدل‌سازی نماید. در سال‌های اخیر از روش SVM در زمینه‌های مختلف هیدرولوژی از جمله پیش‌بینی سطح آب زیرزمینی استفاده شده است [2]. به عنوان نمونه، یون و همکاران (۲۰۱۶) از مدل‌های هوشمند ماشین بردار پشتیبان و شبکه عصبی برای پیش‌بینی سطح آب زیرزمینی در جنوب کره استفاده کردند که نشان دادند مدل ماشین‌بردار پشتیبان دارای کارایی بالاتری از شبکه عصبی مصنوعی در این مطالعات موردی است [3]. اسکندری و همکاران (۲۰۱۶) از ترکیب ماشین‌بردار پشتیبان و تبدیل موجک برای شبیه‌سازی نوسانات سطح آب زیرزمینی دشت برازجان در استان بوشهر در یک دوره ۱۲ ساله استفاده کردند و نشان دادند که مدل ترکیبی نسبت به مدل ماشین‌بردار پشتیبان دقت بالاتری دارد [4].

از دیگر مطالعات انجام شده در شبیه‌سازی سطح آب زیرزمینی در خصوص یادگیری ماشینی SVM به عنوان یک ابزار مدل‌سازی پیشرفته می‌توان به کارهای چیتالپاتی، مالیکارجونا و میلان اشاره کرد [5-7].

حداقل مربعات ماشین‌بردار پشتیبان (LSSVM) به صورت بسطی از SVM مطرح شده است که شامل چندین ویژگی پیشرفته بوده و قابلیت آن توسط محاسبات سریع نشان داده شده است [8]. روش ماشین‌بردار پشتیبان وزن‌دار (WLSSVM) حالت وزن‌دار LSSVM است که با اختصاص دادن وزن به متغیرهای خطا، روش LSSVM بهبود می‌یابد. روش WLSSVM نسبت به LSSVM دارای دقت و عملکرد بهتری است [9].

در روش WLSSVM انتخاب مناسب پارامترهای تنظیم (γ) و کرنل (σ) نقش موثری در بالابردن دقت پیش‌بینی مدل دارند که شناسایی مناسب‌ترین مقادیر

خشکسالی آسیب‌پذیری کمتری دارند. بنابراین آب‌های زیرزمینی به عنوان منبع مهم و با ارزش در تامین آب مورد نیاز بخش‌های مختلف صنعت، کشاورزی و شرب به شمار می‌روند. استفاده بی‌رویه بشر از آب‌های زیرزمینی برای مصارف مختلف سبب شده است که حجم این منابع روز به روز کاهش یابد. استخراج بی‌رویه و خارج از قاعده آب‌های زیرزمینی که معمولاً به مراتب بیش از میزان تغذیه آن است، آثار جانبی زیان‌بار فراوانی از جمله کاهش سطح آب زیرزمینی، خشک شدن چاه‌ها، کاهش آب و یا خشک شدن قنات‌ها، چشمه‌ها و نهرها، تنزل کیفیت آب، افزایش هزینه پمپاژ، نشست زمین و کاهش محصول را در پی خواهد داشت. بنابراین چنانچه منابع آب زیرزمینی به عنوان بخش عظیمی از منابع آب شیرین موجود در کره خاکی به درستی مدیریت نشود، خسارات جبران ناپذیر زیست محیطی، اقتصادی و اجتماعی اجتناب ناپذیر خواهد بود [1].

شبیه‌سازی و پیش‌بینی نوسانات سطح آب زیرزمینی یکی از مهم‌ترین ابزارهای مدیریت و برنامه‌ریزی این منابع با ارزش به شمار می‌رود. مدل‌سازی آب‌های زیرزمینی به عنوان یک ابزار قدرتمند برای کمک به مدیران منابع آب به منظور استفاده بهینه از منابع آب زیرزمینی و حفاظت از این منابع حیاتی به حساب می‌آید. تاکنون مدل‌هایی برای پیش‌بینی سطح آب زیرزمینی ارائه شده است. از جمله این مدل‌ها می‌توان به مدل‌های مبتنی بر روابط فیزیکی و مدل‌های داده‌محور اشاره کرد. اگرچه مدل‌های مبتنی بر روابط فیزیکی ابزاری اساسی برای نشان دادن متغیرهای هیدرولوژیکی و فهم فرآیندهای صورت گرفته در یک سیستم است، ولی این مدل‌ها دارای محدودیت‌هایی هستند. نیاز به مهارت و تخصص برای شبیه‌سازی، لزوم تعیین شرایط مرزی و شرایط اولیه، لزوم شبیه‌بودن فرضیات فیزیکی و ریاضی مدل به سیستم واقعی و هزینه محاسباتی بالا برای شبیه‌سازی دوره‌های بلند مدت از جمله این محدودیت‌ها است. همچنین نبود اطلاعات مربوط به پارامترهای فیزیکی آبخوان این نوع شبیه‌سازی را با محدودیت مواجه می‌کند. این مسئله سبب روی آوردن به مدل‌های داده‌محور به عنوان جایگزین مدل‌های ریاضی و

باشند. شکل (۱) منطقه مورد مطالعه و موقعیت چاه‌ها را در مدل ارتفاعی دشت باغین نشان می‌دهد.

در این مطالعه داده‌های سری زمانی تراز آب زیرزمینی چاه‌های مشاهده‌ای این دشت، از شرکت مدیریت منابع آب تهیه شده است.

مشخصات چاه‌ها در جدول (۱) ارائه شده است. میانگین سطح ایستابی ۱۰ چاه معرف در هنگام ۱۰ سال آماری از ۴۰/۷۱ تا ۱۱۷/۱۲۵ متر متغیر بوده است.

در چاه‌های معرف دشت باغین کمترین تراز آب زیرزمینی مربوط به چاه تلمبه رباطی با مقدار ۱۶۴۷ متر و بیشترین تراز مربوط به چاه گلومک با مقدار ۱۸۴۰/۶۵ است (جدول ۲).

شکل ۱. منطقه مورد مطالعه و موقعیت چاه‌ها در دشت باغین

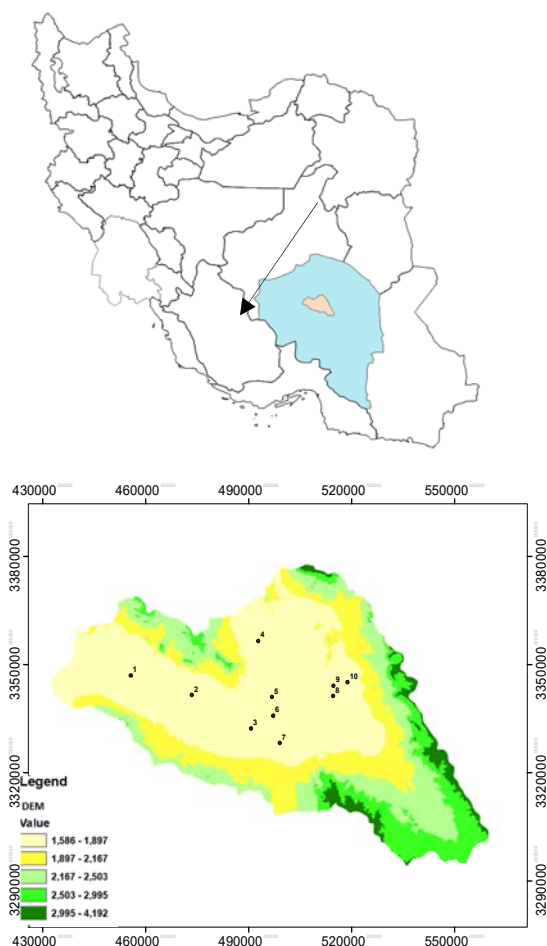


Fig. 1. Study area and wells location in Bagheyn plain.

پارامترهای ذکر شده نیازمند استفاده از الگوریتم‌های بهینه‌سازی است [10, 11]. در مطالعه حاضر از الگوریتم‌های فراکاوشی جامعه پرندگان (PSO) و گرانشی (GSA) برای بهینه کردن پارامترها مورد استفاده قرار می‌گیرد. همچنین انتخاب تابع وزنی مناسب نقش مهمی در بالا بردن دقت پیش‌بینی مدل دارد که در این مطالعه از تابع وزنی دقیق‌تر Myriad استفاده می‌شود [12]. همه برنامه‌ها با نرم‌افزار متلب نوشته شده است. همچنین کارآیی مدل‌های هیبریدی ارائه شده با شاخص‌های آماری ارزیابی شده است.

۲- مواد و روش‌ها

۱-۲- آبخوان مطالعاتی

محدوده مطالعاتی باغین در استان کرمان، جنوبی‌ترین محدوده حوزه آبریز کویرهای درانجیر- ساغند است. این محدوده با وسعت ۵۴۲۰ کیلومتر مربع که ارتفاع متوسط آن ۲۰۹۶ متر بالاتر از سطح دریا است بین طول‌های جغرافیایی ۲۹°-۴۵' تا ۵۷°-۴۰' و عرض‌های جغرافیایی ۳۰°-۳۰' در باند ارتفاعاتی ۱۵۸۶ تا ۴۱۹۲ متر بالاتر از سطح دریای آزاد گسترده است. براساس بررسی‌های بعمل آمده دمای متوسط سالانه ارتفاعات و دشت بر حسب درجه سانتیگراد به ترتیب معادل ۱۲/۸ و ۱۵/۶ است. بطور کلی در دشت باغین سفره تحت فشار و نیمه تحت فشار وجود ندارد. ضخامت آبرفت در این دشت بطور متوسط بین ۱۰۰ تا ۲۵۰ متر متغیر است. دامنه تغییرات ضریب قابلیت انتقال بین ۲۰۰ تا ۳۰۰۰ مترمربع در نقاط مختلف دشت متغیر است، مقدار ضریب ذخیره نیز از کمینه ۲ تا بیشینه ۲۰ درصد در نقاط مختلف دشت متغیر است و ضریب ذخیره متوسط آبخوان ذکر شده در مطالعات اخیر شرکت مدیریت منابع آب ۳ درصد است.

به منظور افزایش دقت در شبیه‌سازی سطح آب زیرزمینی، داده‌های ۱۰ حلقه چاه مشاهده‌ای به صورت ماهانه از مردادماه ۱۳۸۵ تا اسفندماه ۱۳۹۴ مورد استفاده قرار گرفته است. انتخاب چاه‌های مشاهده‌ای مورد نظر به شکلی بوده است که دارای داده‌های کامل تر و قابل اطمینان‌تری نسبت به سایر چاه‌ها

سال ۱۹۹۵ توسط وپنیک و کورینا برای حالت غیرخطی تعمیم یافت که کمترین خطا در طبقه بندی داده‌ها و یا تابع برازش را ارائه داد (14). مبنای کار دسته‌بندی کننده SVM، دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب شود که حاشیه اطمینان بیشتری داشته باشد.

هدف SVM تشخیص تابع $f(x)$ برای الگوهای آموزشی x است، به طوری که بیشترین حاشیه را از مقادیر آموزشی y داشته باشد. به عبارت دیگر، این مدلی است که یک منحنی با ضخامت مشخص به داده‌ها برازش می‌دهد، به طوری که کمترین خطا در داده‌های آموزشی صورت گیرد (15).

مدل SVM به دلیل عدم توانایی در تنظیم دقیق پارامتر هسته و تنها اتکا آن‌ها به بردارهای پشتیبان برای تعیین مرز تصمیم‌گیری، دارای نقاط ضعف است. با وزن دادن به متغیرهای خطا روش LSSVM بهبود یافته است، که بدین ترتیب روش حداقل مربعات ماشین بردار پشتیبان وزن‌دار (WLSSVM) توسط سویکنز و همکاران معرفی شد. مطالعات پیشین نشان داده است تابع‌های پیش‌بینی مدل حداقل مربعات ماشین بردار پشتیبان وزن‌دار قوی‌تر و دقیق‌تر از مدل حداقل مربعات ماشین بردار پشتیبان عمل کرده است [9, 16].

با در نظر گرفتن N عدد مجموعه داده‌های نمونه آموزشی، $x_k \in R^n$ بردار داده‌های ورودی و $y_k \in R^n$ بردار داده‌های خروجی هستند، رگرسیون WLSSVM به عنوان مسئله بهینه‌سازی در فضای وزن‌دار اولیه به صورت زیر فرموله شده است:

$$\min J(w, e) = \frac{1}{2} \|w\|^2 + \frac{1}{2} \gamma \sum_{k=1}^n v_k e_k^2$$

Subject to: (۱)

$$y_k = w^T \phi(x_k) + b + e_k, \quad k = 1, \dots, N$$

که در آن $\phi: R^d \rightarrow R^{\bar{d}}$ اپراتوری است که فضای داده‌های ورودی را به فضاهایی با ابعاد بالاتر نگاشت می‌نماید. بردار $w \in R^{\bar{d}}$ ، بردار وزن در فضای وزن‌دار اولیه است. $e_k = \alpha_k / \gamma$ متغیر خطا، b مقدار بایس و γ پارامتر تنظیم

جدول ۱. محل چاه‌ها

Well ID	Wells name	Wells location	
		Longitude	Latitude
1	Tolombeh-ye Robati	455700	3346900
2	Saadi	473500	3341500
3	Jonob-Bagheyn	490750	3332200
4	Nouzari	492850	3356500
5	Momenabad	496850	3340950
6	Dasht-Bagheyn	497200	3335750
7	Glomak	499100	3328200
8	Tolombe Badi	514650	3341250
9	Dasht-Hosseinabad	514750	3344050
10	Jade Mahan	518870	3345094

Table. 1. Wells location.

۲-۲- روش هیبریدی WLSSVM با الگوریتم‌های فراکاوشی PSO و GSA

در این قسمت در ابتدا اساس ساختار روش WLSSVM، الگوریتم‌های فراکاوشی PSO و GSA توضیح داده می‌شود. سپس چگونگی بهینه کردن پارامترهای موثر روش WLSSVM توسط الگوریتم‌های بهینه سازی فراکاوشی PSO و GSA بیان می‌شود.

جدول ۲. مقادیر آماری سطح آب زیرزمینی در دشت باغین

Well ID	Ground water level (m)	Water table level (m)			
		Min	Max	Mean	SD
1	1647	64.64	81.35	73.55	5.15
2	1686.08	67.21	84	76.83	4.61
3	1774.02	42.66	46.38	44.79	1
4	1740	40.26	41.46	40.94	0.37
5	1759.37	38.78	42.59	40.71	1.15
6	1779.78	53.06	61.94	57.96	2.47
7	1840.65	89.76	93.69	91.68	1.26
8	1781	53.58	59.05	56.11	1.64
9	1774	76.26	79.93	78.32	1.11
10	1780.5	109.2	125.5	117.13	5.04

Table. 2. Statistics value of the groundwater level in the Bagheyn plain.

۲-۲-۱- حداقل مربعات ماشین بردار پشتیبان وزن‌دار

یکی از روش‌های جدید و متداول طبقه‌بندی و تفکیک داده‌ها، ماشین بردار پشتیبان (SVM) است [13]. نخستین بار روش SVM در سال ۱۹۶۳ توسط وپنیک مطرح شد و در

$$\bar{v}_k = \begin{cases} 1 & |e_k / \bar{s}| \leq c_1, \\ \frac{c_2 - |e_k / \bar{s}|}{c_2 - c_1} & c_1 \leq |e_k / \bar{s}| \leq c_2, \\ 10^{-8} & otherwise \end{cases} \quad (11)$$

که در آن $\bar{s}$ تخمین قوی از انحراف معیار متغیرهای خطا $(e_k = \alpha_k / \gamma)$ هستند که به صورت زیر تعریف می‌شود:

$$\bar{s} = \frac{IRQ}{2 \times 0.6745} \quad (12)$$

رنج میانبر IQR، تفاوت بین ۷۵ درصدی و ۲۵ درصدی است. تابع تخمین قوی دیگر انحراف استاندارد $\bar{s} = 1.483MAD(x_i)$ است، که MAD نشانه انحراف مطلق میانه می‌باشد. رابطه (۱۱) به منظور بدست آوردن تخمین قوی هست برای زمانی که توزیع نرمال نباشد. زمانی که پارامترهای (γ, σ) با روش مناسب به درستی تعیین شود e_k دارای یک توزیع متقارن می‌شود. ثابت‌های c_1 و c_2 معمولاً به صورت $c_1 = 2.5$ و $c_2 = 3$ در نظر گرفته می‌شوند. همچنین روش دیگری برای تعیین c_1 و c_2 ، تخمین چگالی توزیع e_k است (9). همچنین روابط محاسبه وزن‌های v_k برای چهار توابع وزنی معمول در مدل WLSSVM به همراه مقدار ضرایب آن‌ها در قسمت‌های بعدی جدول (۵) آورده شده است.

بر اساس تئوری مرسر، تابع کرنل $k(\dots)$ به صورت زیر انتخاب می‌شود:

$$k(x_k, \bar{x}_l) = \langle \varphi(x_k), \varphi(x_l) \rangle_H \quad (13)$$

$$k, l = 1, 2, \dots, n$$

در نتیجه مدل WLSSVM برای تخمین توابع بدین صورت تبدیل می‌شود:

$$y(x) = \sum_{k=1}^N \alpha_k k(x_k, x) + b \quad (14)$$

تابع $k(x_k, \bar{x}_l)$ در واقع یک تابع در فضای اولیه است که

کننده بخش خطا هستند. متغیر خطا (e_k) با فاکتور وزن $\bar{v}_k$ وزن‌دار می‌شود. با توجه به رابطه (۱) و مجموعه داده‌های آموزش، مدل حداقل مربعات ماشین بردار پشتیبان وزن‌دار به صورت زیر تعریف می‌شود:

$$y(x) = w^T \varphi(x) + b \quad (2)$$

به طور کلی ساختار تابع $\varphi(x)$ ناشناخته است. همچنین محاسبه این تابع در فضای مشخصه آن بسیار پیچیده است، از این رو محاسبه غیر مستقیم w از رابطه (۱) غیرممکن است، بنابراین راه حل‌های رگرسیون WLSSVM با ساختن یک تابع لاگرانژی بدست می‌آید:

$$L(w, b, e, x) = J(w, e) - \sum_{k=1}^N \alpha_k \{w^T \varphi(x_k) + b + e_k - y_k\} \quad (3)$$

که در آن α_k ضرایب لاگرانژی است. شرایط بهینه توسط

$$\frac{\partial L}{\partial w} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial e} = 0, \dots, \frac{\partial L}{\partial \alpha_k} = 0 \quad (4)$$

که بعد از حذف w و e نتیجه آن در سیستم کاروس-کوهن-تاکر (KKT) به صورت مجموعه معادلات خطی زیر بدست می‌آید:

$$\begin{bmatrix} \Omega + V_y & 1_n^T \\ 1_n & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix} \quad (5)$$

که در آن:

$$V_y = \text{diag}\{1/\gamma \bar{v}_1, \dots, 1/\gamma \bar{v}_n\} \quad (6)$$

$$\Omega_{k,l} = \langle \varphi(x_k), \varphi(x_l) \rangle_H \quad (7)$$

$$k, l = 1, 2, \dots, n$$

$$y = [y_1, \dots, y_n]^T \quad (8)$$

$$1_n^T = [1, \dots, 1] \quad (9)$$

$$a = [a_1, \dots, a_n] \quad (10)$$

فاکتور وزن $\bar{v}_k$ (Hampel) بدین صورت تخمین زده می‌شود (17, 18).

به صورت ضرب داخلی دو بردار در فضای ویژگی بیان می-شود. به منظور معادل سازی تابع $k(x_k, x_l)$ با ضرب داخلی دو بردار در فضای ویژگی، باید یک تابع $k(x_k, x_l)$ معین مثبت متقارن شود که در شرایط مرسر صدق کند.

ماشین بردار پشتیبان بیشتر با سه تابع کرنل پایه شعاع گوسی RBF، چند جمله ای و تابع کرنل خطی مورد استفاده قرار می گیرد. در روش WLSSVM از تابع پایه شعاع گوسی کرنل $k(x, \bar{x}) = e^{-\frac{\|x_k - x_l\|^2}{\sigma^2}}$ استفاده می شود که در آن σ^2 مقدار ثابت حقیقی و مثبت است که عرض کرنل نامیده می شود. مطالعات پیشین نشان می دهد که تابع کرنل بالاترین دقت را در پیش بینی حاصل می نماید [11].

۲-۲-۲ الگوریتم اجتماع ذرات (PSO)

الگوریتم های فراکاوشی یک نام معمول برای روش هایی هستند که بیشتر برگرفته از طبیعت هستند. در سال های اخیر الگوریتم های فراکاوشی با پیشرفتی سریع در مسائل پیچیده بهینه سازی مورد استفاده قرار گرفته اند. الگوریتم فراکاوشی PSO یک روش فراکاوشی است که در سال ۱۹۹۵ توسط کندی و ابرهات بر اساس رفتار جمعی پرندگان در جستجو و یافتن غذا معرفی شد [19]. در این الگوریتم فرض شده است که پرندگان برای جستجوی غذا به صورت غریزی فاصله خود تا غذا را حس می کنند در حالی که از مکان آن اطلاعی ندارند. علاوه بر این فرض شده است که تمام پرندگان با به اشتراک گذاشتن اطلاعات خود موقعیت نزدیکترین پرنده به غذا را می دانند. در این روش هر جواب مسئله، یک پرنده در فضای جستجو است که ذره نام دارد و به گروه ذرات پرندگان گله گفته می شود. هر ذره دارای یک مقدار شایستگی است که توسط تابع شایستگی مسئله بدست می آید. همچنین هر پرنده دارای یک بردار سرعت است که جهت حرکت پرنده و اندازه سرعت را نشان می دهد. فرآیند حرکت پرندگان به سمت غذا یک پدیده طبیعی است که باعث می شود هر پرنده از گروه پرندگان با سرعت و جهت خاص خود حرکت کند و موقعیت اولیه خود را در گروه بهبود دهد. بردار سرعت بر اساس حافظه

ذخیره شده به وسیله هر پرنده (ذره) بهبود می یابد که این حافظه شبیه حافظه اطلاعات شخصی هر فرد است و از طرف دیگر موقعیت ذره در گروه با توجه به رفتار کلی گروه بهبود می یابد و گروه در مسیر جستجوی خود به مناطق بهتری دست می یابد و این چرخه بهبود موقعیت هر ذره و به دنبال بهبود موقعیت کل گروه ادامه دارد تا گروه به بهترین نقطه (منبع غذایی) برسد. از نظر عددی ذره i ام با موقعیت X و در تکرار $l+1$ ام با رابطه (۱۴) موقعیت خود را بهبود می بخشد که در این فرمول V_i^{l+1} سرعت اصلاح شده ذره i ام است که از رابطه (۱۵) بدست می آید:

$$X_i^{l+1} = X_i^l + V_i^{l+1} \quad (15)$$

$$V_i^{l+1} = \omega V_i^l + c_1 r_1 (pbest_i^l - X_i^l) + c_2 r_2 (gbest_i^l - X_i^l) \quad (16)$$

در این رابطه V_i^l بردار سرعت در تکرار l ام، r_1 و r_2 دو عدد تصادفی بین صفر و یک و $pbest_i^l$ نشان دهنده بهترین موقعیت ذره i ام و $gbest_i^l$ موقعیت ذره در تمام گروه در تکرار l ام است. همچنین c_1 و c_2 پارامترهای اعتماد هستند و تعیین کننده میزان اطمینان به جریان ذرات و یا حرکت گروه هست و ω پارامتر وزن لختی می باشد.

پارامتر ω نقش مهمی را در همگرایی الگوریتم PSO ایفا می کند به شکلی که مقادیر زیاد ω باعث افزایش مقدار بردار سرعت می شود و در هر تکرار گام های بزرگتری برداشته خواهد شد و با کوچک شدن ω مقدار این گام ها کوچکتر می شود که برای رسیدن به جواب بهینه در گام های آخر مفید است، از این رو با قرار دادن یک رابطه پویا به جای مقداری ثابت برای ω به این مهم می توان دست یافت. به علت اهمیت پارامتر وزن در کنترل رفتار الگوریتم PSO در رسیدن به جواب های کلی یا محلی، به جای در نظر گرفتن مقداری ثابت برای ω در تمام طول فرآیند بهینه سازی از رابطه (۱۱) برای محاسبه ω در هر تکرار استفاده می شود تا مقدار ω متناسب با نیاز الگوریتم از یک مقدار ماکزیمم تا مقداری کوچک به صورت خطی تغییر کند. این رابطه توسط شی و ابرهات در سال ۱۹۹۸ پیشنهاد شده است [20].

مقدار اجرام با توجه به برازندگی هر عامل تعیین می‌شوند
[21]:

$$q_i(l) = \frac{fit_i(l) - worst(l)}{best(l) - worst(l)} \quad (20)$$

$$best(l) = \min\{fit_1(t), fit_2(t), \dots, fit_N(t)\} \quad (21)$$

$$wors(l) = \max\{fit_1(t), fit_2(t), \dots, fit_N(t)\} \quad (22)$$

در رابطه بالا $fit_i(l)$ میزان برازندگی جرم i در تکرار l را نشان می‌دهد و $worst(l)$ و $best(l)$ به ترتیب نشان دهنده کمترین و بیشترین میزان برازندگی جرم i در تکرار l است.

$$M_i(l) = \frac{q_i(l)}{\sum_{j=1}^N q_j(l)} \quad (23)$$

در این جا N و $M_i(l)$ به ترتیب تعداد جمعیت و هر جرم عامل i را در تکرار l را نشان می‌دهند.

در این سیستم جرم i با شتابی معادل $a_i^d(l)$ به سمت سایر اجرام در جهت بعد d طبق قوانین جاذبه و دوم نیوتن کشیده می‌شود معادله (۲۴). سرعت هر جرم برابر مجموع ضربی از سرعت فعلی جرم و شتاب جرم است که طبق رابطه (۲۵) تعریف می‌شود. سپس مکان هر جرم عامل با رابطه (۲۶) محاسبه می‌شود:

$$a_i^d(l) = \sum_{j \in kbest, j \neq i} rand_j G(l) \frac{M_j(l)}{R_{ij}(l) + \varepsilon} (x_j^d(l) - x_i^d(l)) \quad (24)$$

$$v_i^d(l+1) = rand_i v_i^d(l) + a_i^d(l) \quad (25)$$

$$x_i^d(l+1) = x_i^d(l) + v_i^d(l+1) \quad (26)$$

که a_i^d ، v_i^d و x_i^d به ترتیب شتاب، سرعت و مکان جرم عامل i در بعد d هستند. $G(l)$ ثابت گرانش، $R_{ij}(l)$ فاصله بین دو جرم i و j و ε یک عدد بسیار کوچک هستند. $rand_i$ و $rand_j$ اعداد تصادفی با توزیع یکنواخت در بازه $[0, 1]$

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{l_{\max}} l \quad (17)$$

که $\omega_{\min}$ و $\omega_{\max}$ به ترتیب مقادیر ماکزیمم و مینیمم وزن لختی است و همچنین $l_{\max}$ تعداد ماکزیمم تکرار است.

الگوریتم PSO دارای کمترین تعداد پارامترهای قابل تنظیم است که همین مسئله استفاده از آن را در عمل بسیار آسان می‌سازد. مقادیر پارامترهای قابل تنظیم در الگوریتم PSO در جدول (۳) ارائه شده است.

جدول ۲. مقادیر پارامترهای PSO برای تخمین سطح آب زیرزمینی

Parameter	Range
Population Size	20
Number of variables	30
Total Iteration Number	500
Omega min , Omega max	0.1 , 0.9
C1 and C2	2

Table. 3. Values of the PSO properties for estimating groundwater level.

۲-۳- الگوریتم جستجوی گرانشی (GSA)

الگوریتم بهینه‌سازی (GSA) اولین بار توسط راشدی و همکاران در سال ۲۰۰۹ با الهام از قانون گرانش در طبیعت و با استفاده از قوانین گرانش نیوتن ارائه شد [21]. در این الگوریتم هر جسم به واسطه نیروی جاذبه، محل و جرم سایر اجسام را درک می‌کند و هر جسم به نسبت میزان جرم خود و فاصله‌ای که با دیگر اجسام دارد، روی سایر اجسام تاثیر می‌گذارد و به آن‌ها نیرو وارد می‌کند. در واقع هر جسم، جسم دیگر را به سمت خود دعوت می‌کند اما تاثیر جسم بزرگتر و نزدیکتر بیشتر است.

در GSA سیستم به صورت مجموعه ای از N جرم تصور می‌شود. هر جرم عامل جوابی از مسئله بهینه سازی است. مکان و سرعت هر جرم عامل i در تکرار l در بعد d به صورت زیر تعریف می‌شود:

$$X_i(l) = \{x_i^1, x_i^2, \dots, x_i^D\} \quad (18)$$

$$V_i(l) = \{v_i^1, v_i^2, \dots, v_i^D\} \quad (19)$$

پژوهش از این تابع وزنی در روش هیبریدی برای رسیدن به دقت بالاتر استفاده شده است.

مقادیر پارامترهای β و δ در تابع‌های وزن Huber و Myriad توسط مدل WLSSVM در نرم‌افزار متلب به صورت خودکار بر اساس بهترین عملکرد انتخاب می‌شود.

جدول ۴. مقادیر توابع وزنی در مدل WLSSVM (12)

Weight Function	v_k	Parameter setting
Huber	$\begin{cases} \beta & e_k/\bar{s} \geq \beta \\ e_k/\bar{s} & e_k/\bar{s} < \beta \end{cases}$	β
Hampel	$\begin{cases} c_2 - e_k/\bar{s} & c_1 \leq e_k/\bar{s} \leq c_2 \\ c_2 - c_1 & \text{other} \end{cases}$	$c_1 = 2.5$
	$\begin{cases} 10^8 & \text{other} \end{cases}$	$c_2 = 3$
Logistic	$\frac{\tanh(e_k/\bar{s})}{(e_k/\bar{s})}$	
Myriad	$\frac{\delta^2}{\delta^2 + (e_k/\bar{s})^2}$	δ

Table 5. Values of the weight functions on WLSSVM model (12).

۲-۴- محاسبه پارامترهای WLSSVM با الگوریتم‌های فراکاوشی PSO و GSA

در مدل WLSSVM، پارامترها تأثیر زیادی در عملکرد آن دارند، بنابراین در انتخاب بهینه پارامترها باید دقت شود. پارامترهای تنظیم (γ) و کرنل (σ) RBF از جمله پارامترهایی هستند که تغییر در آن‌ها باعث تغییر در دیگر پارامترها می‌شود و مقدار بهینه آن‌ها، نقش موثری در بهبود کارایی مدل WLSSVM دارند. از روش‌های مختلف تحلیلی و الگوریتم‌های فراکاوشی می‌توان برای بهبود پارامترهای مدل

هستند که برای حفظ خاصیت تصادفی بودن جستجو استفاده می‌شوند. $kbest$ بیانگر مجموعه k جرم برتر جمعیت است. در تکرارهای اولیه الگوریتم، هنوز مسئله احتیاج به جستجوی مناسب دارد اما با جلو رفتن زمان، جمعیت به نتایج بهتری رسیده است. بنابراین k به صورت تابعی از زمان تعریف می‌شود. به این صورت که در زمان شروع تمام اجرام روی یکدیگر تأثیر می‌گذارند و با گذشت زمان از تعداد اعضا تأثیرگذار بر جمعیت به صورت یک نسبت خطی کم می‌شود تا این که در انتها تنها ۲ درصد از بهترین‌های جمعیت بر سایر اعضا نیرو وارد می‌کنند.

برای تنظیم ثابت گرانش $G(l)$ ، از یک مقدار اولیه شروع کرده و با گذشت زمان مقدار آن کاهش می‌یابد که به صورت زیر تعریف می‌شود:

$$G(l) = G_0 \left(1 - \frac{l}{l_{\max}}\right) \quad (27)$$

که در این رابطه G_0 ثابت گرانش اولیه و $l_{\max}$ ماکزیمم تکرار است. مقادیر پارامترهای قابل تنظیم در الگوریتم GSA در جدول (۴) ارائه شده است.

جدول ۳. مقادیر پارامترهای GSA برای تخمین سطح آب زیرزمینی

Parameter	Range
Population Size	30
Total Iteration Number	500
G_0	30

Table 4. Values of the GSA properties for estimating groundwater level.

۲-۳- تابع‌های وزن در مدل WLSSVM

تابع‌های وزنی مختلفی برای رگرسیون خطی پیشنهاد شده است. چهار تابع مشهور آن به همراه روابط و مقادیر پارامترهای آن در جدول (۵) آورده شده است که سه تابع اول (Huber، Hampel و Logistic) معمول برای رگرسیون است اما تابع چهارم (Myriad) برای محیط‌های سیگنالی پیشنهاد شده است. از آنجا که نوسانات سطح آب زیرزمینی رفتاری مشابه فرآیند سیگنالی دارند، پس بهتر است از این تابع وزنی برای تخمین سطح آب زیرزمینی استفاده شود. همچنین در مطالعات پیشین نشان داده شده است که تابع وزن Myriad دارای بالاترین دقت در میان توابع وزنی در مدل WLSSVM است [12]. بنابراین در این

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (29)$$

$$\text{Subject to } G: \begin{cases} \gamma_{\min} \leq \gamma \leq \gamma_{\max} \\ \sigma_{\min} \leq \sigma \leq \sigma_{\max} \end{cases} \quad (30)$$

که P و O به ترتیب بیانگر مقادیر دقیق (مشاهداتی) و تقریبی (تخمین زده شده) سطح آب چاه هستند. همچنین k تعداد قسمت‌ها در اعتبارسنجی متقاطع و n تعداد داده‌های زیر مجموعه آزمایشی هستند. مقادیر تخمین زده سطح آب چاه (O_i) بستگی به انتخاب مناسب پارامترهای مدل WLSSVM دارند.

$\gamma_{\min}$ و $\gamma_{\max}$ حد پایین و بالای پارامتر تنظیم مدل WLSSVM است که به ترتیب برابر با 10^{-5} و 10^0 هستند. همچنین $\sigma_{\min}$ و $\sigma_{\max}$ حد پایین و بالای پارامتر کرنل RBF مدل WLSSVM است که آن‌ها نیز به ترتیب برابر با 10^{-5} و 10^0 هستند [22].

در این مطالعه داده‌های ورودی تراز سطح ایستابی در هنگام ۱۰ سال آماری (۱۳۹۴-۱۳۸۵) است که مجموعاً ۱۲۰ داده هستند. از روش معمول اعتبارسنجی متقابل k-fold برای کارایی مدل در انتخاب مجموعه داده‌ها استفاده شده است. در اعتبارسنجی متقابل تعداد تقسیم‌ها، معمولاً متناسب با تعداد داده‌ها هستند. در این مطالعه از ۷۵٪ داده‌ها به عنوان زیر مجموعه آموزش و ۲۵٪ برای زیر مجموعه آزمایش استفاده شده است. در مجموع با استفاده از اعتبارسنجی متقاطع k-fold با $k=4$ ، تعداد ۹۰ داده که به عنوان بخش آموزش و ۳۰ داده برای آزمایش استفاده شده است. بنابراین از یک زیر مجموعه برای فرآیند آزمایش و سه زیر مجموعه دیگر برای فرآیند آموزش استفاده شده است (شکل ۴).

۲-۵- کارایی مدل

روش حداقل مربعات ماشین بردار پشتیبان وزن‌دار به منظور پیش‌بینی سطح آب چاه‌ها استفاده شده است. ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد برای آزمایش مدل استفاده شده است.

WLSSVM استفاده کرد. در این مطالعه از الگوریتم‌های فراکاوشی بهینه‌سازی PSO و GSA در تعیین مقادیر بهینه پارامترهای مدل WLSSVM استفاده شده است. چگونگی ترکیب روش‌های WLSSVM با الگوریتم‌های فراکاوشی PSO و GSA به صورت جداگانه در شکل‌های (۲ و ۳) نشان داده شده است.

شکل ۲. فلوچارت PSO برای یافتن پارامترهای بهینه WLSSVM

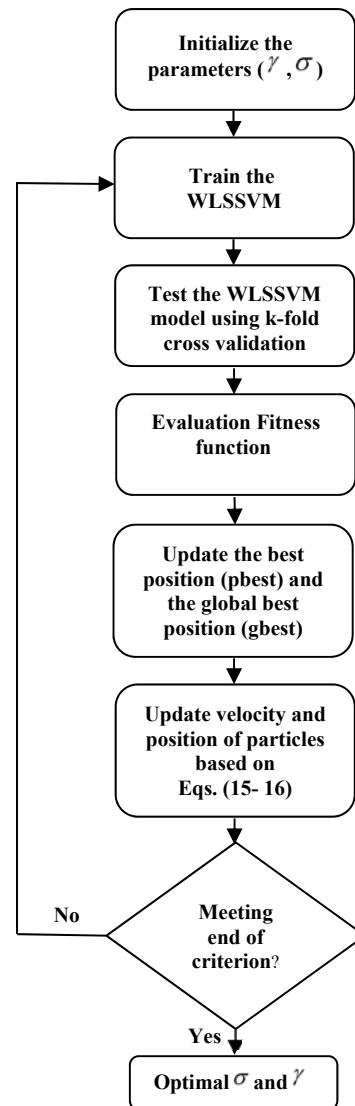


Fig. 2. Flowchart of PSO for finding optimal WLSSVM parameters

تابع هدف الگوریتم‌های فراکاوشی PSO و GSA برای شناسایی مقادیر بهینه پارامترهای تنظیم (γ) و کرنل

$$\text{Minimise } \text{Fitness} = \sum_{i=1}^k \frac{RMSE_i}{k} \quad (28)$$

به RBF(σ) است:

شکل ۳. فلوچارت GSA برای یافتن پارامترهای بهینه WLSSVM

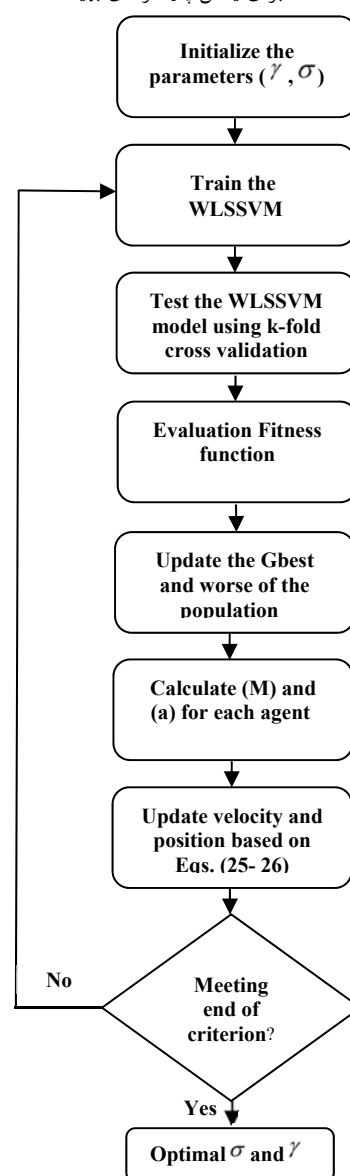


Fig. 3. Flowchart of GSA for finding optimal WLSSVM parameters.

پارامترهای آماری میانگین خطای مطلق (MAE)، جذر میانگین مربعات خطا (RMSE)، درصد میانگین خطای مطلق (AAE) و راندمان مدل (EF) برای داده‌های آزمایشی به وسیله روابط زیر محاسبه می‌شوند (23):

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - O_i| \quad (31)$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2 \right]^{\frac{1}{2}} \quad (32)$$

$$AAE = \left[\frac{\sum_{i=1}^n \frac{(O_i - P_i)}{O_i}}{n} \right] \times 100 \quad (33)$$

$$EF = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (\bar{O}_i - O_i)^2} \quad (34)$$

که O و P به ترتیب نمادی برای مقادیر دقیق (مشاهداتی) و تقریبی (تخمین زده شده) سطح آب چاه هستند. همچنین تعداد داده‌های آزمایشی است که در این مطالعه ۳۰ است. مقادیر نسبتاً کم و نزدیک به صفر برای ضرایب MAE، RMSE، AAE و مقدار نزدیک به ۱ ضریب EF حاکی از دقت بالا در نتایجی است که مدل ارائه شده تخمین زده است.

۳- نتایج و بحث

۳-۱- توزیع خطای گوسین در مدل هیبریدی WLSSVM-GSA

برای ده چاه معرف توزیع خطا گوسین در مدل هیبریدی WLSSVM-GSA بدست آمده است که به عنوان نمونه مقدار توزیع و هیستوگرام خطای نمونه چاه نوذری در شکل (۵) نشان داده شده است. همانگونه که پیشتر ذکر شد تعداد داده آموزشی ۹۰ است.

شکل ۴. چهار قسمت داده‌های آموزش و آزمایش

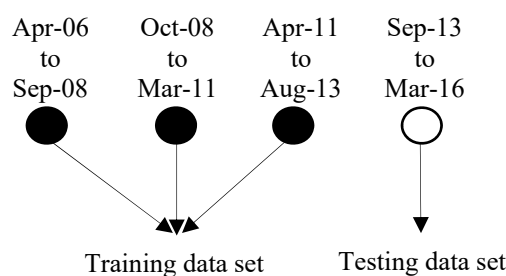


Fig. 4. Four-fold training and testing datasets.

به منظور بررسی عملکرد مدل‌های هیبریدی، شاخص‌های آماری برای ۱۰ چاه محاسبه شده است که در جدول‌های (۷ و ۸) ارائه شده است. در هر دوی مدل هیبریدی WLSSVM-GSA و WLSSVM-PSO مقادیر ضرایب AAE، RMSE، AAE نزدیک به صفر و مقدار ضریب EF بسیار نزدیک به ۱ برای داده‌های آزمایشی هستند که نشان دهنده کارایی مدل‌ها هستند.

مدل زمانی دقیق هست که EF آن بیشتر از ۰/۸ باشد (24)، در مدل‌های هیبریدی ارائه شده در ۱۰ چاه معرف همگی بالای ۰/۹۹۶ بدست آمده است که این قابلیت و توانمندی مدل‌های هیبریدی را نشان می‌دهد.

در مدل‌های هیبریدی کمترین مقادیر RMSE، MAE، AAE مربوط به چاه نوذری و بیشترین این ضرایب مربوط به چاه جاده ماهان می‌باشد. همچنین بیشترین EF مربوط به چاه نوذری و کمترین این ضریب مربوط به چاه جاده ماهان است. شکل (۷) نشان می‌دهد در همه چاه‌های معرف مقدار راندمان مدل در روش WLSSVM-GSA نسبت به روش WLSSVM-PSO بیشتر است که نشان دهنده کارایی بیشتر مدل WLSSVM-GSA هست.

شکل ۶. نتایج راندمان مدل سطح آب زیرزمینی در چاه‌های مشاهداتی با روش‌های WLSSVM-PSO و WLSSVM-GSA

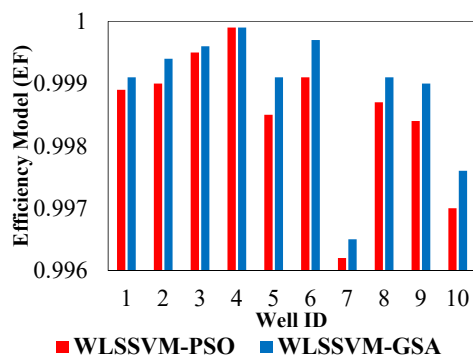


Fig. 6. Efficiency model (EF) on the resulting groundwater level at the observed wells on WLSSVM-PSO and WLSSVM-GSA models.

شکل ۵. مقدار و هیستوگرام توزیع خطای گوسین e_k در چاه نوذری

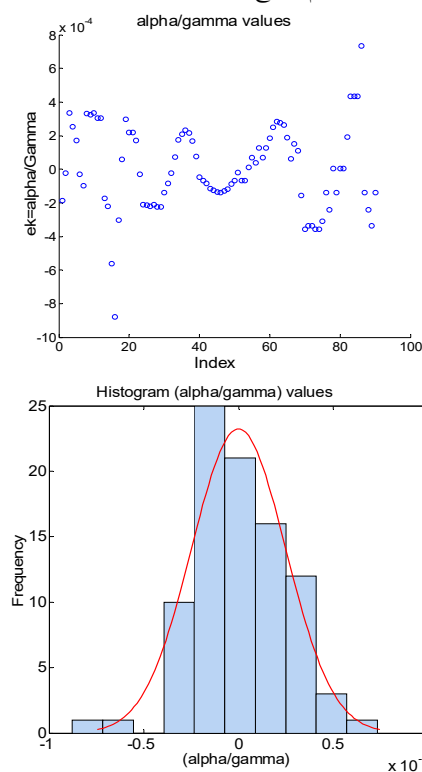


Fig. 5. Values and histogram for distribution of e_k for Nouzari well.

۳-۲- مقایسه نتایج ارزیابی مدل هیبریدی فراکاوشی

جدول (۶) مقادیر بهینه پارامترهای تنظیم (γ) و کرنل (σ) RBF توسط الگوریتم های فراکاوشی در روش WLSSVM را نشان می‌دهد. مقادیر پارامتر تنظیم (γ) و کرنل (σ) RBF در روش هیبریدی WLSSVM-GSA نسبت به روش WLSSVM-PSO بیشتر است.

جدول ۵. مقادیر بهینه پارامترهای WLSSVM بدست آمده از PSO و GSA در ۱۰ چاه مشاهداتی.

Well ID	WLSSVM Parameters			
	WLSSVM-PSO		WLSSVM-GSA	
	σ	γ	σ	γ
1	519.44	8084.90	552.66	9139.39
2	433.12	4572.20	618.51	9197.39
3	338.27	9776.40	427.41	9431.96
4	528.40	11018.95	690.13	13959.0
5	442.48	6041.28	528.25	9268.92
6	491.52	4860.22	612.48	8336.47
7	455.76	6299.20	591.54	11205.77
8	550.67	7482.02	716.05	12079.13
9	332.91	3241.09	487.43	6948.17
10	243.63	2889.54	341.08	5643.96

Table. 6. Optimal WLSSVM parameters achieved PSO and GSA for observed wells.

شکل ۷. مقایسه نتایج مشاهداتی و تخمینی سطح آب زیرزمینی با دو

روش WLSSVM-GSA و WLSSVM-PSO در چاه جاده ماهان

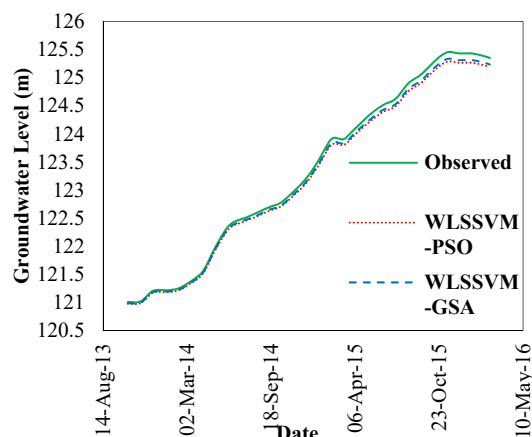


Fig. 7. Comparison of the observed groundwater level and the WLSSVM-PSO and WLSSVM-GSA estimated groundwater level for Jade-Mahan well.

۴- نتیجه گیری

در این پژوهش از مدلی هیبریدی فراکاوشی ماشین بردار پشتیبان وزن دار برای تخمین سری زمانی آب زیرزمینی استفاده شده است. مدل WLSSVM خود به تنهایی دارای دقت خوبی برای تخمین توابع دارد اما از آنجا که پارامترهای تنظیم (γ) و کرنل (σ) RBF نقش موثری در دقت این مدل داشته است، برای یافتن مقادیر بهینه این پارامترها از الگوریتم‌های بهینه‌سازی فراکاوشی PSO و GSA استفاده شده است. از تابع معمول کرنل RBF در روش WLSSVM و دقیق‌ترین تابع وزنی Myriad نیز در این مدل‌ها استفاده شده است. به منظور بررسی کارایی این روش، داده‌های ۱۰ حلقه چاه در دشت باغین استان کرمان در هنگام دوره زمانی ۱۰ ساله استفاده شد. در این مطالعه مجموعه داده‌ها، با استفاده از اعتبارسنجی متقاطع k-fold ($k=4$) برای آموزش و آزمایش تقسیم شده است. در مجموع با استفاده از اعتبارسنجی متقاطع k-fold با $k=4$ ، تعداد ۹۰ داده که به عنوان بخش آموزش و ۳۰ داده برای آزمایش استفاده شده است. میزان خطای داده‌های ۱۰ چاه معرف همه گوسین بدست آمده است. دقت و اثربخشی این روش جدید هیبریدی با استفاده از شاخص‌های آماری بررسی شد که مقادیر بسیار قابل قبولی بودند. مقایسه داده‌های مشاهداتی و تخمین زده هر دو روش هیبریدی نشان داد که برابری بسیار نزدیکی بین این داده‌ها است. اما روش هیبریدی WLSSVM-GSA دقت بالاتری نسبت به روش هیبریدی WLSSVM-PSO

جدول ۶. مقادیر نتایج شاخص‌های آماری در چاه‌ها با روش

WLSSVM-PSO

Well ID	Statistical indicators by using WLSSVM-PSO model			
	MAE	RMSE	AAE	EF
1	0.0276	0.0293	0.0343	0.9989
2	0.0294	0.0316	0.0356	0.9990
3	0.0039	0.0043	0.0084	0.9995
4	0.0002	0.0003	0.0004	0.9999
5	0.0076	0.0082	0.0181	0.9985
6	0.0176	0.0192	0.0287	0.9991
7	0.0189	0.0199	0.0202	0.9962
8	0.0195	0.0217	0.0332	0.9987
9	0.0077	0.0081	0.0097	0.9984
10	0.0944	0.106	0.0760	0.9970

Table. 7. Results of statistical indicators on WLSSVM-PSO model.

جدول ۷. مقادیر نتایج پارامترهای آماری در چاه‌ها با روش

WLSSVM-GSA

Well ID	Statistical indicators by using WLSSVM-GSA model			
	MAE	RMSE	AAE	EF
1	0.0259	0.0276	0.0323	0.9991
2	0.0207	0.0223	0.0251	0.9994
3	0.0035	0.0039	0.0076	0.9996
4	0.00003	0.00005	0.00002	0.9999
5	0.0060	0.0065	0.0143	0.9991
6	0.0103	0.0113	0.0168	0.9997
7	0.0151	0.0158	0.0161	0.9965
8	0.0139	0.0156	0.0237	0.9991
9	0.0053	0.0055	0.0066	0.9990
10	0.0676	0.0759	0.0544	0.9976

Table. 8. Results of statistical indicators on WLSSVM-GSA model.

۳-۳- برابری نتایج حاصل از مدل هیبریدی

فراکاوشی با مقادیر مشاهداتی

برای چاه جاده ماهان که دارای کمترین دقت از مابین ۱۰ چاه معرف بود، تخمین سطح آب زیرزمینی با مدل‌های هیبریدی WLSSVM-GSA و WLSSVM-GSA در شکل (۷) نشان داده شده است. در هر دو روش از تابع وزنی Myriad استفاده شده است. برابری نمودارهای مقادیر مشاهداتی و تخمین زده سطح ایستابی با دو روش، حاکی از کارایی و دقت بالای روش هیبریدی فراکاوشی ماشین بردار پشتیبان وزن دار است.

- parameters based on a robust system architecture. *Neural Computing and Applications*. 2018; 30(8): 2355-69.
- [12] De Brabanter K, Pelckmans K, De Brabanter J, Debruyne M, Suykens JA, Hubert M, et al., editors. Robustness of kernel based regression: a comparison of iterative weighting schemes. *International conference on artificial neural networks*; 2009: Springer.
- [13] Deng N, Tian Y, Zhang C. Support vector machines: optimization based theory, algorithms, and extensions: CRC press; 2012.
- [14] Vapnik V. The nature of statistical learning theory: Springer science & business media; 2013.
- [15] Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995; 20(3): 273-97.
- [16] Quan T, Liu X, Liu Q. Weighted least squares support vector machine local region method for nonlinear time series prediction. *Applied Soft Computing*. 2010;10(2):562-6.
- [17] David H. Early sample measures of variability. *Statistical Science*. 1998:368-77.
- [18] Rousseeuw PJ, Leroy AM. Robust regression and outlier detection: John Wiley & sons; 2005.
- [19] Kennedy J, Eberhart R, Shi Y. *Swarm Intelligence* Morgan Kaufmann Publishers. San Francisco. 2001.
- [20] Shi Y, Eberhart R, editors. A modified particle swarm optimizer. 1998 IEEE international conference on evolutionary computation proceedings IEEE world congress on computational intelligence (Cat No 98TH8360); 1998: IEEE.
- [21] Rashedi E, Nezamabadi-Pour H, Saryazdi S. GSA: a gravitational search algorithm. *Information sciences*. 2009;179(13):2232-48.
- [22] Hoang N-D, Pham A-D. Hybrid artificial intelligence approach based on metaheuristic and machine learning for slope stability assessment: A multinational data analysis. *Expert Systems with Applications*. 2016;46:60-8.
- [23] Kalantari M, Akbarpour A, Khatibinia M. Numerical Modeling of Groundwater Flow in Unconfined Aquifer in Steady State with Isogeometric Method. 2018.
- [24] Lal A, Datta B. Development and implementation of support vector machine regression surrogate models for predicting groundwater pumping-induced saltwater intrusion into coastal aquifers. *Water Resources Management*. 2018; 32(7): 2405-19.
- PSO دارد. به دلیل انعطاف پذیری بالای مدل‌های هوشمند پیشنهاد می‌شود که مدل هیبریدی فراکاوشی ماشین بردار پشتیبان وزن دار برای تخمین سطح آب چاه انجام شود.
- ### ۵- منابع
- [1] Zinivand A. Groundwater Level Modeling with Combining Artificial Neural Network and Wavelet (Case Study: Sharif Abad Plain). MSc Thesis, University of Qom, Qom. 2014.
- [2] Deka PC. Support vector machine applications in the field of hydrology: a review. *Applied soft computing*. 2014; 19: 372-86.
- [3] Yoon H, Hyun Y, Ha K, Lee K-K, Kim G-B. A method to improve the stability and accuracy of ANN-and SVM-based time series models for long-term groundwater level predictions. *Computers & geosciences*. 2016; 90: 144-55.
- [4] Eskandari A, Solgi A, Zarei H. Simulating Fluctuations of Groundwater Level Using a Combination of Support Vector Machine and Wavelet Transform. *Irrigation Sciences and Engineering Journal*. 2016; 41(1):165-80.
- [5] Ch S, Mathur S. Groundwater level forecasting using SVM-PSO. *International Journal of Hydrology Science and Technology*. 2012; 2(2):202-18.
- [6] Mallikarjuna B, Sathish K, Krishna PV, Viswanathan R. The effective SVM-based binary prediction of ground water table. *Evolutionary Intelligence*. 2020:1-9.
- [7] Ghourdoyee Milan S, Aryaazar N, Javadi S, Razdar B. Simulation of groundwater head using LS-SVM and comparison with ANN & MLR. *Hydrogeology*. 2020;5(1):118-33.
- [8] Suykens JA, Van Gestel T, De Brabanter J. Least squares support vector machines: World scientific; 2002.
- [9] Suykens JA, De Brabanter J, Lukas L, Vandewalle J. Weighted least squares support vector machines: robustness and sparse approximation. *Neurocomputing*. 2002;48(1-4):85-105.
- [10] Aljarah I, Ala'M A-Z, Faris H, Hassonah MA, Mirjalili S, Saadeh H. Simultaneous feature selection and support vector machine optimization using the grasshopper optimization algorithm. *Cognitive Computation*. 2018; 10(3): 478-95.
- [11] Faris H, Hassonah MA, Ala'M A-Z, Mirjalili S, Aljarah I. A multi-verse optimizer approach for feature selection and optimizing SVM

Hybrid Learning Machine Metaheuristic Model for Estimating Groundwater Level

Shiva Khosravi¹, Amir Robati^{2*}

1. Department of Civil Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran.
2. Department of Civil Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran

*A_robati@iauk.ac.ir

Abstract

Groundwater is the most reliable source of supply for potable water and supports a wide array of economic and environmental services. There is a significant concern that groundwater levels are declining due to intense aquifer use. The sustainable management of groundwater resources requires good planning and concerted efforts. To manage groundwater resources, it is necessary to predict the groundwater levels and its fluctuations. The prediction groundwater level can guide water managers and engineers effectively. On the other hand, there are multifarious types of equipment for measuring levels of groundwater. Sophisticated water level loggers or divers can measure the groundwater level automatically. Sounding devices with acoustic and light signals are also used to check groundwater levels. The use of devices for measuring the level of groundwater is time-consuming and costly. To reduce the time and cost of the groundwater level measuring process, many methods of Artificial Intelligence (AI) have been utilized for estimating the groundwater level. Among the AI methods, SVMs has great ability in predicting non-linear hydrological processes. Support vector machines (SVMs) is as an intelligent computational method for predicting hydrological processes. Recently, (SVMs) have been successfully applied in classification problems, regression and predicting; as techniques of machine learning, statistics and mathematical analysis. The SVM is based on the structural risk minimization (SRM), which can escape from various difficulties, such as the necessity of a large number of control parameters and a local minimum in artificial neural networks (ANNs). The weighted least squares support vector machines (WLSSVM) was first introduced by Suykens et al., and has proved to be much more robust in several fields, especially for noise mixed data, than least squares version of SVM (LSSVM). Their powerful scientific research provides motivation for employing WLSSVM method in estimating groundwater level. The accurate value of WLSSVM parameters (γ, σ) effect on the estimation, these optimal parameters can be achieved optimization algorithms. Therefore, weighted least square support vector machine (WLS-SVM) model was coupled with particle swarm optimization (PSO) and gravitational search algorithm (GSA) as metaheuristic algorithms for estimating well water level. In this study, an attempt has been made to use the hybrid model with high accuracy to estimate the groundwater level. In order to estimate the groundwater level, ten wells data in Bagheyne plain of Kerman province is considered during ten-year time series. The estimated value obtained by the WLSSVM-PSO and WLSSVM-GSA models are compared with the observed value, and showed the estimated results have nearly coincidence with observed values. Numerical results show the merits of the suggested technique for groundwater level simulation. In order to verify the hybrid learning machine metaheuristic model, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Average Absolute Error (AAE), and Model Efficiency (EF) are computed, and these statistical indicators stand on the good acceptable range, and find WLSSVM-GSA is more accurate than WLSSVM-PSO. The results demonstrate that the new hybrid WLSSVM-GSA model has high efficiency and accuracy with observed values, and the modelling method is an innovative and powerful idea in estimating well water level.

Keywords: Groundwater level – Weighted least square support vector machine – Particle swarm optimization – Gravitational search algorithm – Efficiency.